FEBS Letters 413 (1997) 282–288 FEBS 19068

NMR characterization of the full-length recombinant murine prion protein, *m*PrP(23–231)

Roland Riek, Simone Hornemann, Gerhard Wider, Rudi Glockshuber, Kurt Wüthrich*

Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule-Hönggerberg, CH-8093 Zürich, Switzerland

Received 7 July 1997

Abstract The recombinant murine prion protein, *m*PrP(23–231), was expressed in *E. coli* with uniform ^{15}N-labeling. NMR experiments showed that the previously determined globular three-dimensional structure of the C-terminal domain *m*PrP(121–231) is preserved in the intact protein, and that the N-terminal polypeptide segment 23–120 is flexibly disordered. This structural information is based on nearly complete sequence-specific assignments for the backbone amide nitrogens, amide protons and α-protons of the polypeptide segment of residues 121–231 in *m*PrP(23–231). Coincidence of corresponding sequential and medium-range nuclear Overhauser effects (NOE) showed that the helical secondary structures previously identified in *m*PrP(121–231) are also present in *m*PrP(23–231), and near-identity of corresponding amide nitrogen and amide proton chemical shifts indicates that the three-dimensional fold of *m*PrP(121–231) is also preserved in the intact protein. The linewidths in heteronuclear ^{1}H–^{15}N correlation spectra and ^{15}N{^{1}H}-NOEs showed that the well structured residues 126–230 have correlation times of several nanoseconds, as is typical for small globular proteins, whereas correlation times shorter than 1 nanosecond were observed for all residues of *m*PrP(23–231) outside of this domain.

Key words: Prion protein; NMR; Protein structure and dynamics; Transmissible spongiform encephalophaties; Correlation time

1. Introduction

The 'protein-only' hypothesis [1,2] claims that transmissible spongiform encephalopathies (TSE) are distinct from infectious processes caused by bacteria, viruses or viroids in that nucleic acids are apparently not essential for the propagation of the infectious agent [3], which has been termed 'prion' [4]. TSEs have been linked with a conformational polymorphism of the 'prion protein' (PrP) [5], where a disease-related transformation of the ubiquitous cellular form of the protein, PrP^{C}, into the infectious scrapie form, PrP^{Sc}, is believed to consist of a change from a predominantly α-helical protein to a β-sheet-containing fold [6,7]. Considering the central role thus attributed to PrP, investigations of the three-dimensional structure of this protein are of keen interest. We previously solved the NMR structure of a self-folding C-terminal domain of the cellular form of the murine prion protein, *m*PrP(121–231) [8–10]. Here, we report a preliminary structural characterization of the intact polypeptide chain of mature murine PrP, *m*PrP(23–231), in aqueous solution.

*Corresponding authors: R.G. [Fax: (41) (1) 633-1036] and K.W. [Fax: (41) (1) 633-1151].

Besides the NMR structure determination of *m*PrP(121–231) [9], discussions on PrP conformations have so far been founded primarily on optical spectroscopy. PrP^{C} consists of a single polypeptide chain that contains two glycosylation sites, and is attached to the cell surface by a glycosyl-phosphatidylinositol anchor at its carboxy-terminus [11]. After separation from the cell membrane, it is a water-soluble, protease K-sensitive protein for which circular dichroism (CD) spectroscopy indicates a high content of helical secondary structure [12,13]. PrP^{Sc} has so far only been observed as an insoluble oligomer that displays resistance to protease K digestion and has characteristics of an amyloid [6,7]. Based on Fourier transform reflection infrared spectroscopy it was concluded that a significant percentage of the polypeptide chain in PrP^{Sc} forms β-sheet secondary structure [12,13]. Collection of more detailed structural data is all the more important as no one has succeeded so far to generate infectious PrP^{Sc} in vitro, either from previously denatured infectious material, or from recombinant or synthetic PrP or fragments thereof (see also ref. [14]). Three-dimensional structure determinations of both functional forms of PrP, and possibly of refolding intermediates arising in the course of the disease-causing conformational transition, promise to contribute in essential ways to a rational basis for continued investigations on the role of PrP and possibly additional factors in the pathology of TSEs, such as the Creutzfeldt-Jakob disease (CJD) in humans and bovine spongiform encephalopathy (BSE) in cattle.

2. Materials and methods

NMR experiments were carried out either with uniformly ^{15}N-labeled murine prion protein, *m*PrP(23–231), that was prepared as described in the preceding paper [15], or with the ^{15}N-labeled C-terminal domain *m*PrP(121–231) [8,9]. The protein concentration was 0.8 mM in a solvent of 90% H_2O/10% D_2O containing 0.01 mM EDTA, 0.01 mM PMSF, 0.1 μM Pepstatin and the protease inhibitor cocktail COMPLETE (Boehringer-Mannheim) at pH = 4.5 and T = 293 K. NMR spectra were recorded on a 750 MHz Varian UNITYplus spectrometer and on a 500 MHz Bruker DRX spectrometer. Two-dimensional (2D) [^{15}N,^{1}H]-COSY spectra [16] were acquired using spectral widths of ω_1 = 1800 Hz and ω_2 = 10 000 Hz. For *m*PrP(121–231) the maximal evolution times where t_{1max} = 45 ms and t_{2max} = 204 ms, and the time domain data size was 160×4096 points. For *m*PrP(23–231) we used t_{1max} = 36 ms, t_{2max} = 204 ms and a time domain data size of 128×4096 points. ^{15}N{^{1}H}-NOEs of *m*PrP(23–231) were measured using the procedure described by Dayie and Wagner [17] with a relaxation delay of 3 s and a proton saturation length of 3 s achieved by applying a train of 120 degree pulses in 10 ms intervals. The dataset was recorded with ω_1 = 2440 Hz, ω_2 = 7000 Hz, t_{1max} = 33 ms, t_{2max} = 37 ms and a time domain data size of 160×512 points. Three-dimensional (3D) ^{15}N-resolved [^{1}H,^{1}H]-NOESY spectra [18] were measured with ω_1 = 1800 Hz, ω_2 = 10 000 Hz, ω_3 = 10 000 Hz, t_{1max} = 14 ms, t_{2max} = 23 ms, t_{3max} = 102 ms, a mixing time τ_m = 40 ms and a time domain data size of 50×460×2048 points. Data process-

PII S0014-5793(97)00920-4

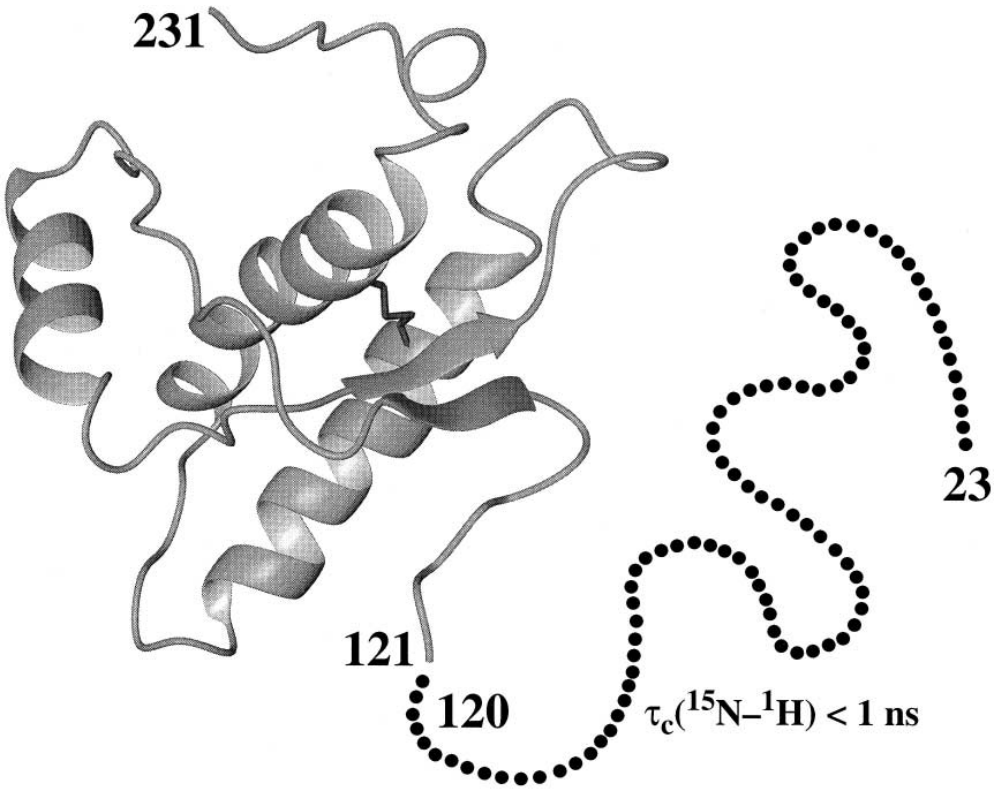

Fig. 1. Survey of the global structural characterization of *m*PrP(23–231) achieved in this paper. The fold that was previously observed in the isolated C-terminal domain, *m*PrP(121–231) [9], is preserved in the intact protein, and the backbone ^{15}N–^1H moieties of the well-structured residues 126–230 manifest rotational correlation times of several ns, which is typical for a folded protein of this size. Dots represent the 98 residues of the N-terminal segment 23–120, which shows features of a flexible, 'random coil-like' polypeptide, with rotational correlation times for the ^{15}N–^1H groups of $\tau_c < 1$ ns.

ing was performed with the program PROSA [19], and the spectral analysis was supported by the program XEASY [20].

3. Results and discussion

For ease of presentation Fig. 1 visualizes the key conclusions from the work reported in this paper: the previously determined three-dimensional fold of the C-terminal domain *m*PrP(121–231) is preserved in the intact protein. The backbone ^{15}N–^1H moieties in this domain have effective rotational correlation times, τ_c, of several nanoseconds, which is typical for small globular proteins [21]. The N-terminal polypeptide segment 23–120 has a small dispersion of the proton chemical shifts and τ_c-values < 1 ns for the ^{15}N–^1H moieties, which is typical for a flexible 'random coil-like' polypeptide chain. In the following we describe the NMR data that lead to this global structural characterization of *m*PrP(23–231).

The NMR-spectral analysis in this paper is largely based on the previously obtained, nearly complete sequence-specific resonance assignments for the isolated C-terminal domain *m*PrP(121–231) [9] and on sequence-specific assignments for the backbone ^{15}N, $^1H^N$ and $^1H^\alpha$ atoms of the residues 121 to 231 in intact *m*PrP(23–231). The assignments for *m*PrP(23–231) were based on sequential NOE connectivities [22–24], using 2D [^{15}N,^1H]-COSY and 3D ^{15}N-resolved [^1H,^1H]-NOESY. The assignments are complete with the following exceptions: The Xxx-Pro connectivities in positions 136/137 and 157/158, the segment 165–170 and Phe^{176}, which could be only partially assigned also in *m*PrP(121–231) [9], and the segment 220–223 (Fig. 2).

Comparison of corresponding ^{15}N and $^1H^N$ chemical shifts for the residues 121–231 in *m*PrP(121–231) and *m*PrP(23–231) suggests that the three-dimensional structure of this polypeptide segment is very similar in the two proteins. The deviations of the chemical shifts in *m*PrP(121–231) from the random coil shifts (Fig. 3a and b) are typical for a globular protein and indicate that even minor conformational rearrangements would be manifested by readily measurable shifts [24]. The shift differences between the two proteins (Fig. 3c and d), however, are very small. The ensemble of the chemical shift data in Fig. 3 thus provides strong evidence that the three-dimensional fold of the structurally well-defined polypeptide segment 126–230 in *m*PrP(121–231) [9] is preserved in the intact *m*PrP(23–231). In addition, in *m*PrP(23–231) three helices and a single helical turn near the C-terminus could be identified from the patterns of d_{NN}, $d_{\alpha N}(i,i+3)$ and $d_{\alpha N}(i,i+4)$ NOE connectivities (Fig. 2) obtained from a 3D ^{15}N-resolved [^1H,^1H]-NOESY spectrum recorded with a mixing time of 40 ms [25]. These regular secondary structures are in exactly corresponding sequence locations to those of the helices that were previously identified in *m*PrP(121–231) [9]. A long-range NOE Tyr$^{163}H^N$–H^NMet^{129} identified in *m*PrP(23–231) has a counterpart in the antiparallel β-sheet of *m*PrP(121–231) [9], and thus supports the presence of a corresponding β-sheet in the intact protein, which is independently implicated by the chemical shifts.

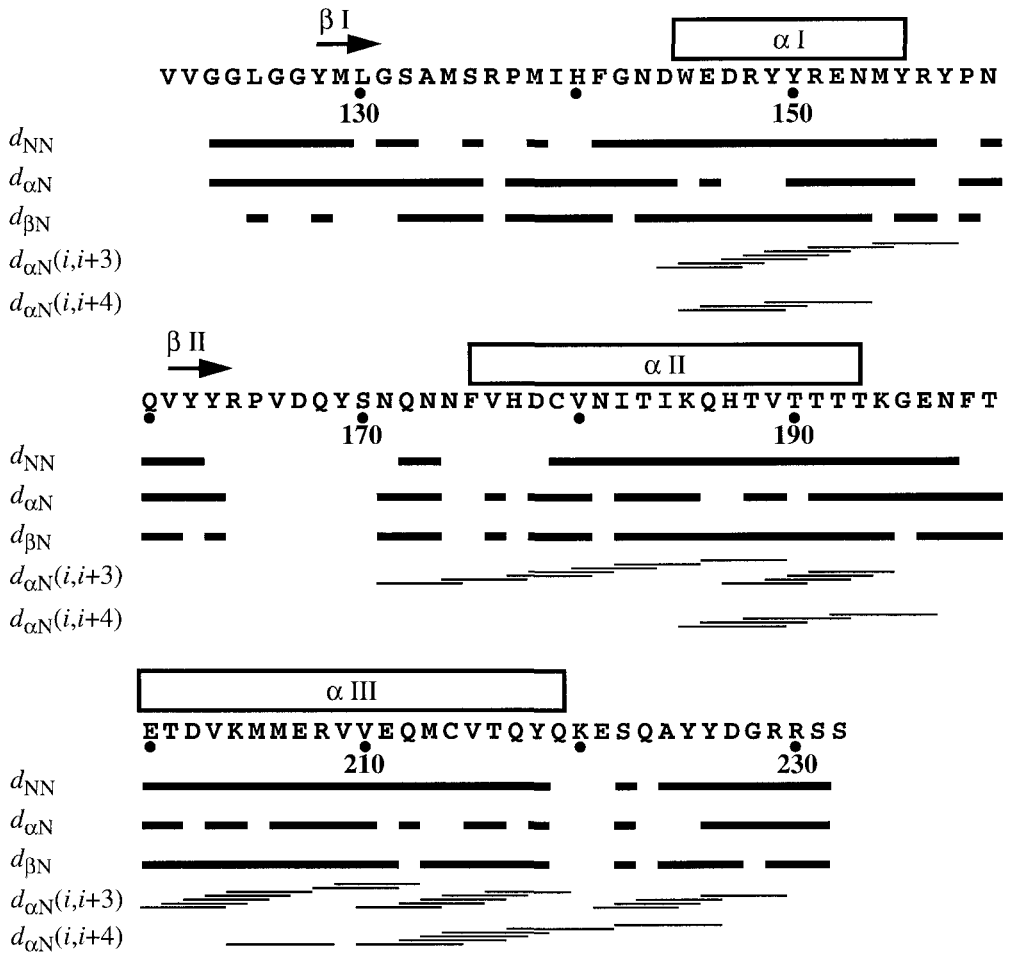

Fig. 2. Amino acid sequence of the polypeptide segment 121–231 and survey of the sequential and medium-range NOE connectivities observed in *m*PrP(23–231). The sequential NOE connectivities d_{NN}, $d_{\alpha N}$ and $d_{\beta N}$ are indicated with black bars, and the medium-range connectivities $d_{\alpha N}(i,i+3)$ and $d_{\alpha N}(i,i+4)$ are shown by lines starting and ending at the positions of the residues that are related by the NOE. The sequence locations of the regular secondary structure elements identified in the isolated C-terminal domain *m*PrP(121–231) are shown above the sequence, with α for helices and β for β-strands.

The structural characterization of *m*PrP(23–231) as summarized in Fig. 1 is also supported by NMR data that relate to the segmental flexibility of the polypeptide chain, i.e. the linewidth of the [^{15}N,^{1}H]-COSY cross peaks, and the sign and magnitude of the ^{15}N{^{1}H}-NOEs [24]. In the [^{15}N,^{1}H]-COSY spectrum of *m*PrP(23–231) (Fig. 4a) two qualitatively different kinds of peaks can be distinguished, i.e, a first group of about 100 broad peaks with full linewidths at half height along ω_2(^{1}H) of about 25 Hz, and a second group of sharp, more intense peaks with linewidths of about 12 Hz (Fig. 5a). Comparison with the [^{15}N,^{1}H]-COSY spectrum of *m*PrP(121–231) (Fig. 4b) showed that the positions of the broad peaks in the intact protein coincide very closely with the peaks in the free domain, where the linewidths are about 20 Hz. On the basis of the sequence-specific assignments (Fig. 2) all the broad peaks in *m*PrP(23–231) could be attributed to residues of the polypeptide segment 126–230, and 6 sharp peaks were attributed to residues 122–125, 231 and 232 (Ser^{232} is an additional residue in the construct used to express the protein). A count of the remaining sharp peaks revealed approximately 50 peaks in the spectral area outside of the two rectangles in Fig. 4a, and of the order of 25 Gly peaks inside the solid rectangle. Thus, there are about 75 sharp peaks that can be attributed as a group to the total number of 86 non-proline amide groups in the polypeptide segment 23–121, and there is not a single broad peak that could be attributed to the residues 23–121.

^{15}N{^{1}H}-NOEs were recorded using a 2D[^{15}N,^{1}H]-COSY-relayed ^{15}N{^{1}H}-NOE experiment recorded at a ^{1}H frequency of 500 MHz [17]. Here, one expects peaks with positive sign and relative intensity +1 for ^{15}N–^{1}H groups with effective rotational correlation times $\tau_c \gg 1$ ns, peaks with negative sign and relative intensity −4 for $\tau_c \ll 1$ ns, and peaks with very small positive or negative intensities for τ_c-values near 1 ns. These three limiting situations are illustrated in Fig. 5b, where Met^{129}, Ile^{139}, Phe^{141} and Val^{180} show positive NOEs, Val^{122} is almost nulled, and there are three strong negative NOEs that were attributed to ^{15}N–^{1}H^{N} moieties in the segment 23–121. In Fig. 4c and d, the subspectra of negative and positive ^{15}N{^{1}H}-NOEs have been separated. Fig. 4d mimics the spectrum of broad [^{15}N,^{1}H]-COSY peaks that were as-

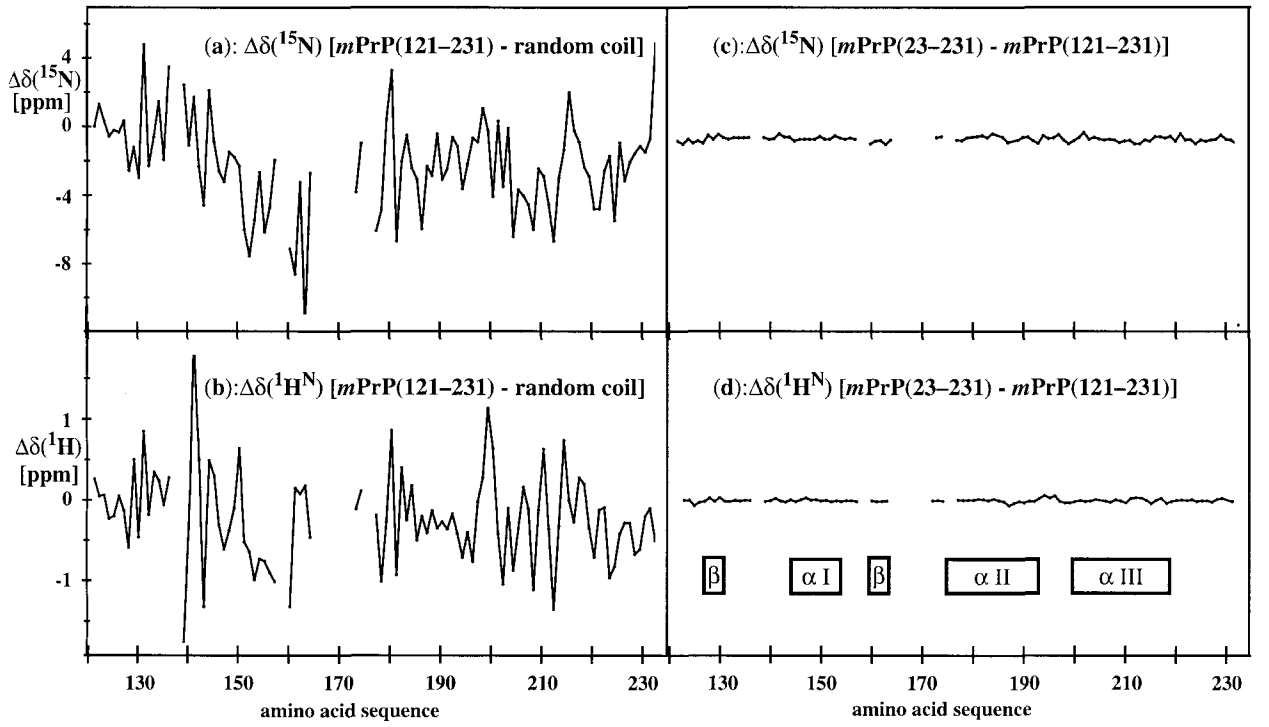

Fig. 3. Plots of chemical shift differences for the amide ^{15}N and $^{1}H^{N}$ resonances of the residues 121–231 in *m*PrP versus the amino acid sequence. (a) and (b): Difference between the experimental shifts for *m*PrP(121–231) and the 'random coil' shifts for ^{15}N [29] and $^{1}H^{N}$ [24], respectively. (c) and (d): Difference between corresponding chemical shifts of ^{15}N and $^{1}H^{N}$, respectively in *m*PrP(121–231) and *m*PrP(23–231). The sequence locations of the regular secondary structures of *m*PrP(121–231) are indicated in (d), with α for helices and β for β-strands.

signed to the polypeptide segment 126–230, thus confirming that the residues in the C-terminal domain are restrained in their segmental flexibility. In Fig. 4c nearly all of the approximately 75 sharp [^{15}N,^{1}H]-COSY peaks attributed to residues 23–121 are also represented by a negative $^{15}N\{^{1}H\}$-NOE, thus confirming that this polypeptide segment is highly flexible.

In addition to the backbone amide groups the Trp indole $^{15}N^{\varepsilon}$–^{1}H moieties were also investigated to characterize the side chain mobility of these residues (Figs. 4 and 6). In the solution structure of *m*PrP(121–231) the ring of Trp^{145} is located on the surface and is solvent exposed [9]. Nonetheless, the $^{15}N\{^{1}H\}$-NOE data reflect slow rotational tumbling, with $\tau_c \gg 1$ ns. In contrast, the envelope of the resonances from the indole $^{15}N\{^{1}H\}$-NOEs of the 5 Trp residues 57, 65, 73, 81 and 89 in the octapeptide repeats, Trp^{31} and Trp^{99} has a negative sign and thus indicates that most, or possibly all of the Trp side chains in the segment 23–120 of *m*PrP(23–231) have high mobility, with $\tau_c < 1$ ns. Note that these 7 Trp residues are distributed in a regular fashion over the polypeptide segment 30–100.

4. Conclusions and outlook

The structural characterization summarized in Fig. 1 is to be attributed to the cellular form PrP^{C}, of the mouse prion protein, since it is based on experimental data collected in aqueous solution without addition of chemical denaturants or detergents. The present study answers a key question that has been raised (e.g. [6,26,27]) following the NMR structure determination of *m*PrP(121–231) [9]: The absence of the N-terminal segment of residues 23–120 in *m*PrP(121–231) does not significantly affect the C-terminal domain, which has the same three-dimensional fold in the full-length protein (Fig. 1). This is an important result because the polypeptide segment 90–145 in mammalian prion proteins has been reported to have polymorphic traits in structure predictions as well as in experimental studies such as, for example, structure characterization in different solvents [26].

The presence of an extensive polypeptide segment with the properties of a flexible extended coil in PrP^{C} could enable structural transitions to PrP^{Sc} aggregates that might display sizeable β-sheet content [6,7,28] without major conformational rearrangements in the C-terminal domain 121–231. In particular, the combination of the results presented in this paper with earlier findings that the polypeptide segment 90–120 is also protected against protease K digestion in PrP^{Sc} [6,7] emphasizes that there is a major change in the structural arrangement of the residues 90–120 in PrP^{C} and PrP^{Sc}.

The global view of the *m*PrP(23–231) structure in Fig. 1 leaves room for future refinements. For the C-terminal globular domain, extension of the sequence-specific backbone assignments to the amino acid side chains will provide a platform for studies of possible effects of the N-terminal chain extension on subtle features of the tertiary structure. For the N-terminal flexible coil of residues 23–120, sequence-specific assignments will enable a search for possible variations in the degree of flexibility along the sequence. For example, since only about 75 ^{15}N–$^{1}H^{N}$ peaks have so far been attributed to this molecular region of 86 ^{15}N–$^{1}H^{N}$ moieties, it will be of interest whether the remainder of the resonances might have escaped detection because of some slow, local exchange processes in this overall highly flexible polypeptide segment. This

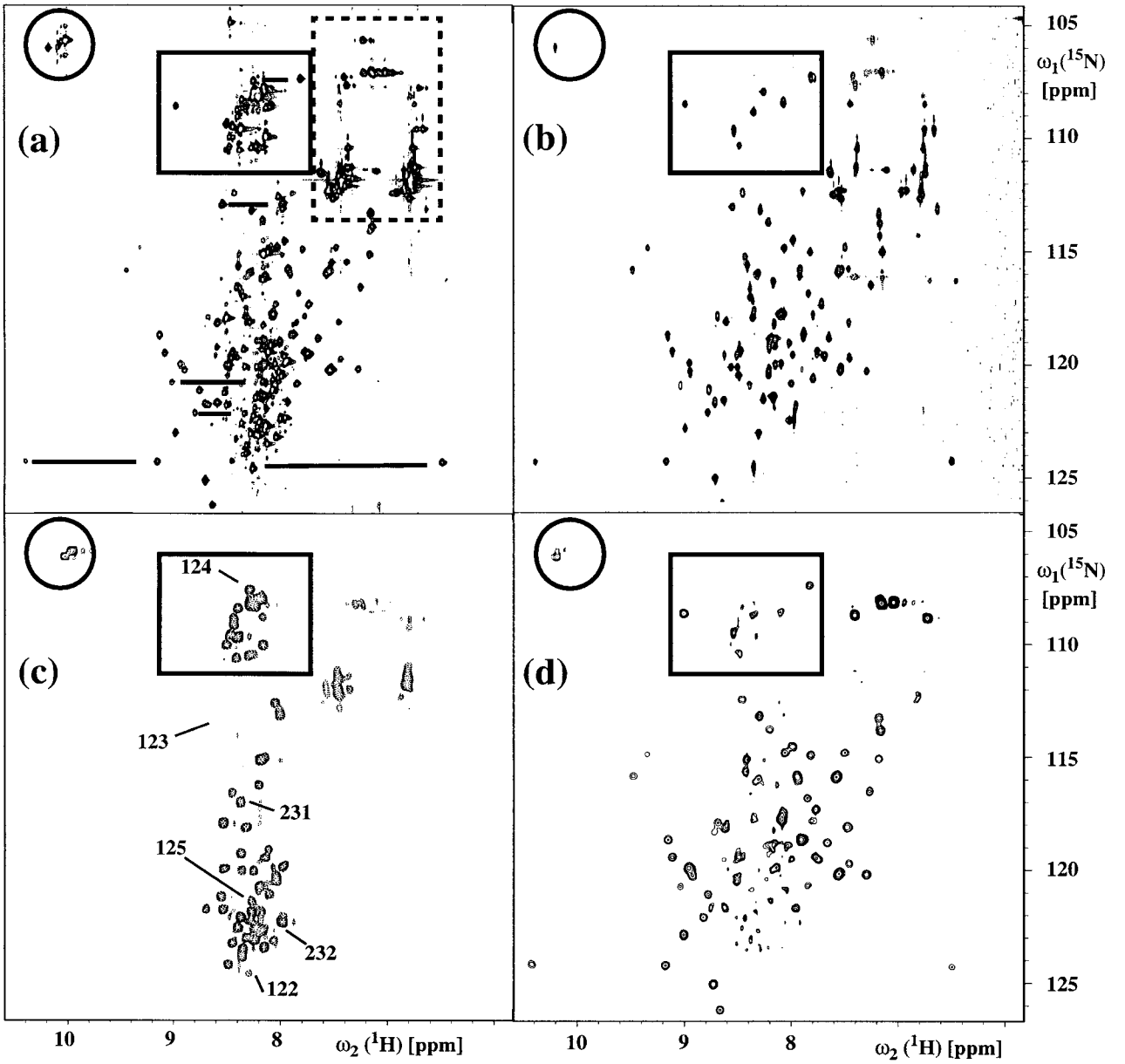

Fig. 4. (a) [^{15}N,^{1}H]-COSY spectrum of *m*PrP(23–231). The horizontal bars indicate the locations of the cross sections shown in Fig. 5a. The rectangle drawn with broken lines surrounds the spectral region of the side chain amide and guanidino resonances. (b) [^{15}N,^{1}H]-COSY spectrum of *m*PrP(121–231). (c) Subspectrum containing the negative peaks from a [^{15}N[^{1}H]-COSY-relayed ^{15}N{^{1}H}-NOE experiment with *m*PrP(23–231). The peaks that were individually assigned to residues in the polypeptide segment 121–231 are identified with the residue number. (d) Subspectrum containing the positive peaks from the same experiment as (c). In all four spectra the circle contains the Trp indole $^{15}N^{\varepsilon}$–^{1}H cross peaks, and the rectangle the Gly backbone ^{15}N–$^1H^N$ peaks.

might, among others, provide new insight into the structural and functional roles of the Pro-containing octapeptide repeats in positions 51–91 of *m*PrP(23–231).

Acknowledgements: Financial support was obtained from the Schweizerischer Nationalfonds (Projects 31.49047.96, 438+.050285 and 438+.050287). S.H. is supported by a grant from the Boehringer Ingelheim Fonds. We thank Mrs. R. Hug for the careful processing of the manuscript.

References

[1] Alper, T., Cramp, W.A., Haig, D.A. and Clarke, M.C. (1967) Nature 214, 764–766.
[2] Griffith, J.S. (1967) Nature 215, 1043–1044.
[3] Riesner, D. (1996) Chemie Unserer Zeit 30, 66–74.
[4] Prusiner, S.B. (1982) Science 216, 136–144.
[5] Büeler, H., Aguzzi, A., Sailer, A., Greiner, R., Autenried, M., Aguet, M. and Weissmann, C. (1993) Cell 73, 1339–1347.
[6] Prusiner, S.B. (1996) Trends Biochem. Sci. 21, 482–487.
[7] Weissmann, C. (1996) FEBS Lett. 389, 3–11.
[8] Hornemann, S. and Glockhuber, R. (1996) J. Mol. Biol. 262, 614–619.
[9] Riek, R., Hornemann, S., Wider, G., Billeter, M., Glockshuber, R. and Wüthrich, K. (1996) Nature 382, 180–182.
[10] Billeter, M., Riek, R., Wider, G., Hornemann, S., Glockshuber, R. and Wüthrich, K. (1997) Proc. Nat. Acad. Sci. USA 94, in press.
[11] Stahl, N., Baldwin, M.A., Teplow, D.B., Hood, L., Gibson,

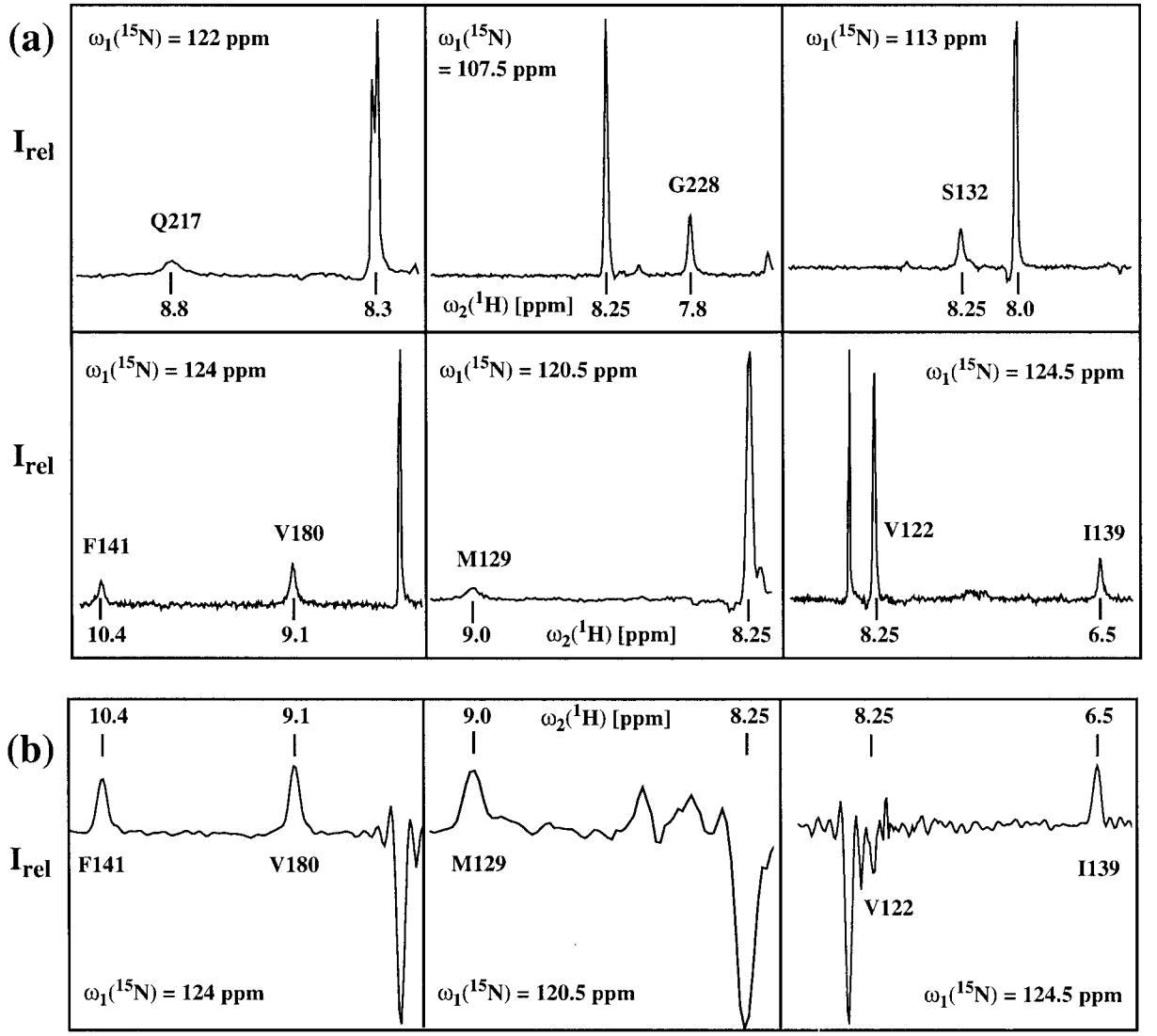

Fig. 5. (a) Cross sections along the $\omega_2(^1H)$ frequency taken at the positions indicated with horizontal bars in the [^{15}N,^1H]-COSY spectrum of Fig. 4a. To allow for the fact that the $\omega_1(^{15}N)$ positions of the individual peaks contained in a given cross section may be slightly different, cross sections taken exactly on-resonance for each individual peak in the cross section were added to obtain the correct representation of the relative peak intensities. (b) Cross sections through the [^{15}N,^1H]-COSY-relayed $^{15}N\{^1H\}$-NOE spectrum of Fig. 4, c+d. The resonances belonging to residues in the assigned segment 121–231 of *m*PrP(23–231) are identified by residue type and sequence position.

B.W., Burlingame, A.L. and Prusiner, S.B. (1993) Biochemistry 32, 1991–2002.
[12] Caughey, B.W., Dong, A., Bhat, K.S., Ernst, D., Hayes, S.F. and Caughey, W.S. (1991) Biochemistry 30, 7672–7680.
[13] Pan, K.M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I., Huang, Z., Fletterick, R.J., Cohen, F.E. and Prusiner, S.B. (1993) Proc. Natl. Acad. Sci. USA 90, 10962–10966.
[14] Glockshuber, R., Hornemann, S., Riek, R., Wider, G., Billeter, M. and Wüthrich, K. (1997) NATO Advanced Research Workshop on Prions and Brain Diseases in Animals and Humans, Erice, Italy, August 19–22, 1996, in press.
[15] Hornemann, S., Korth, C., Oesch, B., Riek, R., Wider, G., Wüthrich, K. and Glockshuber, R. (1997) FEBS Lett. 19096.
[16] Bodenhausen, G. and Ruben, D.J. (1980) Chem. Phys. Lett. 69, 185–191.
[17] Dayie, K.T. and Wagner, G. (1994) J. Magn. Reson. A 111, 121–126.
[18] Fesik, S.W. and Zuiderweg, E.R.P. (1988) J. Magn. Reson. 78, 588–593.
[19] Güntert, P., Dötsch, V., Wider, G. and Wüthrich, K. (1992) J. Biomol. NMR 2, 619–629.
[20] Bartels, C., Xia, T., Billeter, M., Güntert, P. and Wüthrich, K. (1995) J. Biomol. NMR 6, 1–10.
[21] Richarz, R., Nagayama, K. and Wüthrich, K. (1980) Biochemistry 19, 5189–5196.
[22] Billeter, M., Braun, W. and Wüthrich, K. (1982) J. Mol. Biol. 155, 321–346.
[23] Wagner, G. and Wüthrich, K. (1982) J. Mol. Biol. 155, 347–366.
[24] Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley, New York.
[25] Wüthrich, K., Billeter, M. and Braun, W. (1984) J. Mol. Biol. 180, 715–740.
[26] Harrison, P.M., Bamborough, P., Daggett, V., Prusiner, S.B. and Cohen, F.E. (1997) Curr. Opin. Struct. Biol. 7, 53–59.
[27] Wickner, R.B. (1997) Prion Diseases of Mammals and Yeast: Molecular Mechanisms and Genetic Features, p. 49, Springer, New York, Berlin.
[28] Horwich, A.L. and Weissmann, J.S. (1997) Cell 89, 499–510.
[29] Braun, D., Wider, G. and Wüthrich, K. (1994) J. Am. Chem. Soc. 116, 8466–8469.

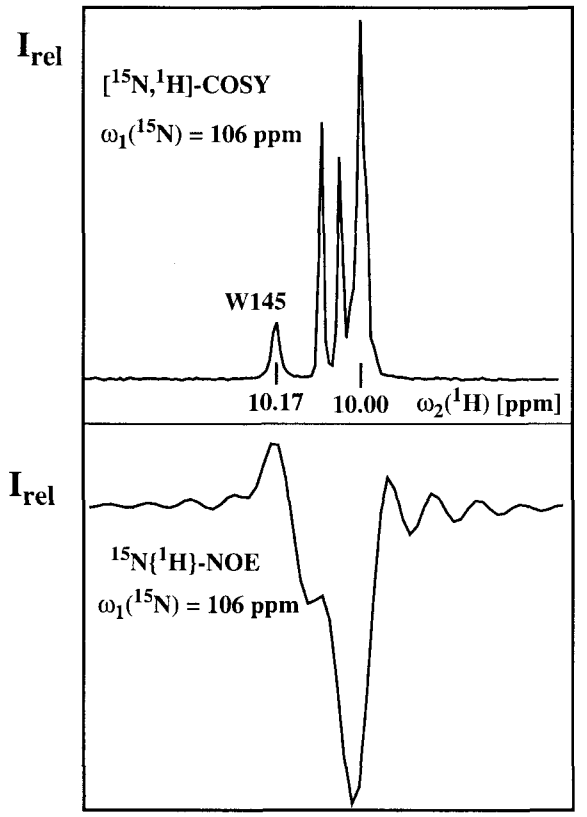

Fig. 6. Cross sections through the spectra (a) and (c+d) in Fig. 4, showing the resonances of the indole $^{15}N^{\varepsilon}$–^{1}H moieties of the eight tryptophan residues in *m*PrP(23–231). The cross sections were obtained as described in the caption to Fig. 5. The peak of the single Trp residue from the assigned polypeptide segment 121–231 is identified in the top trace.